ESD-TR-77-187

Semiannual Technical Summary

Information Processing Techniques Program

Volume II: Communications-Adaptive Internetting

31 March 1977

Prepared for the Defense Advanced Research Projects Agency
under Electronic Systems Division Contract F19628-76-C-0002 by

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. This work was sponsored by the Defense Advanced Research Projects Agency under Air Force Contract F19628-76-C-0002 (ARPA Order 2929).

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Raymond L. Loiselle

Raymond L. Loiselle, Lt. Col., USAF
Chief, ESD Lincoln Laboratory Project Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

INFORMATION PROCESSING TECHNIQUES PROGRAM
VOLUME II: COMMUNICATIONS-ADAPTIVE INTERNETTING

SEMIANNUAL TECHNICAL SUMMARY REPORT
TO THE
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

1 JULY 1976 – 31 MARCH 1977

ISSUED 3 AUGUST 1977

LEXINGTON MASSACHUSETTS

ABSTRACT

This report describes work performed on the Communications-Adaptive Internetting program sponsored by the Information Processing Techniques Office of the Defense Advanced Research Projects Agency. The Semi-annual Technical Summary preceding this one was entitled "Airborne Command and Control" and covered the period ending 30 June 1976. In order to come into compliance with the current fiscal year schedule, this report has been delayed three months and covers the period 1 July 1976 through 31 March 1977.

CONTENTS

INFORMATION PROCESSING TECHNIQUES PROGRAM

COMMUNICATIONS-ADAPTIVE INTERNETTING

I. INTRODUCTION AND SUMMARY

The goal of this program is the investigation and development of techniques for communications-adaptive internetting with particular emphasis on digital voice communications. The specific class of problems addressed relates to packet speech networks whose interconnecting links may be stressed as a result of traffic overloads or are time varying due to natural or hostile actions. This program extends the technology of fixed-topology packet-switching speech communications networks by introducing techniques for adapting both source encoder algorithms and network utilization strategies to the time-varying character of the links in an attempt to provide sustained voice communications capability under adverse conditions.

The major accomplishments of the past nine months are summarized below.

A. Variable-Data-Rate Protocols for C^2 Links

The LOS-UHF variable-rate protocol has been modified to comply with a newly developed modem capability which eliminates the need for transmission of high-rate probe signals on the channel. This new protocol yields a significant decrease in upward switching delay. The broadcast satellite channel model described in the previous Semiannual Technical Summary (SATS) has been defined in more detail and implemented on the PDP-11/45.

B. Demonstration System for Packetized Voice over a Single-User Link with Time-Varying Capacity

A demonstration system in which the packetized output of one variable-rate speech encoder is transmitted through a simulated radio channel with time-varying capacity has been developed and is currently being implemented on the PDP-11/45 and several peripheral processors. The demonstration facility will operate in real time and will allow a person to speak through the system while the capacity of the channel changes due to the addition of varying amounts of Gaussian noise to the channel signals (simulating jamming, fading, etc.). The system responds to capacity changes by appropriately changing the modem data rate and speech encoding bit rate. The system includes a Lincoln Digital Signal Processor (LDSP) to run a set of vocoder algorithms, a flexible signal conditioner to allow computer-controlled switching of sampling clocks and filters, and a pair of Lincoln Digital Voice Terminals (LDVTs) to implement the transmitter and receiver of a variable-rate modem.

C. Adaptive Variable-Rate Modem with "No-Cost" Up-Probing

A modem design which eliminates the need for explicit up-probing for channel assessment purposes has been developed. The technique involves a comparison between signal detection decisions made by integrating over the current duration T_b of the channel signals, with decisions made over smaller intervals $T_b/2$, $T_b/4$, etc. If decisions over smaller intervals match those over T_b, the possibility of higher rate operation is indicated. The notion supports a variety of signaling strategies. An MFSK implementation has been chosen for implementation on the LDVT in anticipation of computational efficiency through the use of the FFT.

D. Adaptive Voice Communication on an Integrated, Time-Varying Network

A networking strategy appropriate for handling variable-rate packet speech communication in a large, integrated network has been designed. The strategy combines the Packetized Virtual Circuit (PVC) concept,[1] an embedded coding technique of speech compression, and a specially tailored variable-rate communication protocol. A dual-rate adaptive predictive coding (APC) speech algorithm, which is well matched to this networking strategy, has been developed. A simple model has been set forth for the traffic generated by many voice users sharing a single wideband link, and analysis of this model has led to a simple expression for the fraction of packets which must be discarded due to user competition for such a link.

II. VARIABLE-DATA-RATE COMMUNICATION PROTOCOLS FOR C^2 LINKS

A. Modified LOS Variable-Rate Protocol

The LOS-UHF variable-rate protocol described in the last SATS included periodic transmission of upward-probing signals in order to detect possible increases in channel capacity. Since that time, a variable-rate modem has been developed which allows up-probing without explicit transmission of high-rate probe signals on the channel. For each received baud, the receive modem produces an assessment of the highest channel rate that could have been sustained over this baud. The method for making these assessments will be described in Sec. IV. The receiving nodal processor uses these baud-by-baud rate assessments to make rate-changing decisions.

To comply with this new modem capability, the protocol of the LOS channel has been revised, implemented, and observed on the dynamic link display. The receiver records the minimum of the rate assessments over 1024 bits (256 bauds) and sends a probe reply just as it did using the original protocol. The probe reply contains a rate change command to one-half the minimum of the rate assessments. The rate switching delay to an increase in capacity when channel loading = 0 is

$$T_{UD} = \underbrace{\frac{3}{2\mu_F C_{AB}}}_{\substack{\text{average time}\\ \text{to assess new}\\ \text{capacity}}} + \underbrace{\frac{3}{2\mu_C C_{BA}}}_{\substack{\text{average time}\\ \text{for reply to}\\ \text{reach remote}\\ \text{node}}} + \underbrace{\frac{1}{2\mu_C C_{AB}}}_{\substack{\text{average time}\\ \text{to complete}\\ \text{transmission}\\ \text{of a control}\\ \text{packet in}\\ \text{progress}}} + \underbrace{\frac{1}{\mu_C C_{AB}}}_{\substack{\text{time to send}\\ \text{rate-hike}\\ \text{message}}}$$

where

$1/\mu_F$ = full packet size in bits

$1/\mu_C$ = control packet size in bits

C_{AB} = original channel capacity (bps)

C_{BA} = capacity of return link (bps).

Since the receiver is continuously making this assessment every 1024 bits, the upward switching delay is greatly reduced from the case analyzed in the previous SATS, where periodic transmission of probe signals was used. Figure 1 plots upward rate switching delays for the new protocol.

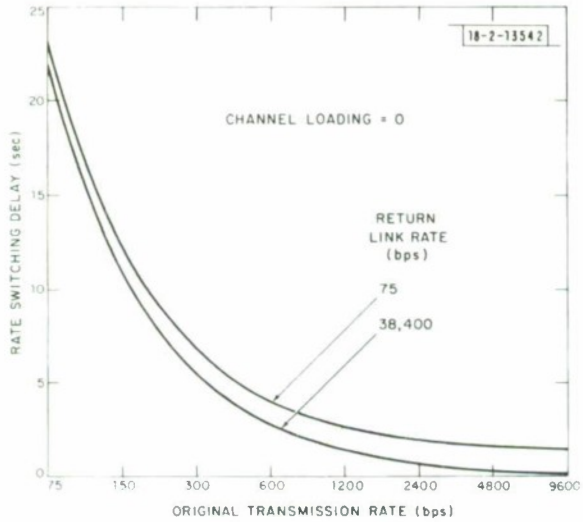

Fig. 1. Average-rate switching delay to capacity increase relative to original transmission rate and rate of return link of modified LOS channel.

B. Broadcast Satellite Variable-Rate Protocol Implementation

The broadcast satellite channel model and protocol described in the previous SATS has been developed further and implemented on the PDP-11/45. The model consists of a satellite repeater and four ground transceivers as shown in Fig. 2. The presence of uplink jamming alone has been considered in designing the channel protocol; therefore, the link capacities change simultaneously, and the data rates are uniform. The model employs rates of 600, 1200, 2400, and 4800 bps, probing to 9600 bps.

The communication format is shown in Fig. 3. A frame consists of eight 1024-bit slots, the first of which is dedicated to the transmission of reservation packets by the four transmitters, and the remainder of which are message slots. Reservations are made for the frame following the current frame due to the 0.25-sec round-trip propagation delay as shown in Fig. 4. During a frame in which no message slots have been reserved, a message packet may be transmitted on a random-access basis.

Flow charts of the transmitter and receiver protocols are shown in Figs. 5 and 6. To detect an increase in capacity, an upward probe is transmitted in the last message slot of every other frame by alternating transmitters. Each of the receivers will prepare to send a reply if the assessed capacity is greater than the current rate. A reply may accompany a reservation packet or a message packet. Upon receiving a probe reply, all receivers cancel any pending reply requests. At the beginning of the frame following that in which the reply was received, all transmitters will change rate and request their receivers to do the same at the appropriate time. Upward switching delay is given by

$$T_{UD} = \underbrace{\frac{2n-1}{2\mu_F C}}_{\text{average time until next probe}} + \underbrace{\frac{1}{\mu_F C}}_{\text{probe time}} + \underbrace{\frac{n}{\mu_F C}}_{\text{time of next frame during which reply will be received}}$$

where

$1/\mu_F$ = full packet size in bits

C = original channel capacity (bps)

n = number of full packets per frame.

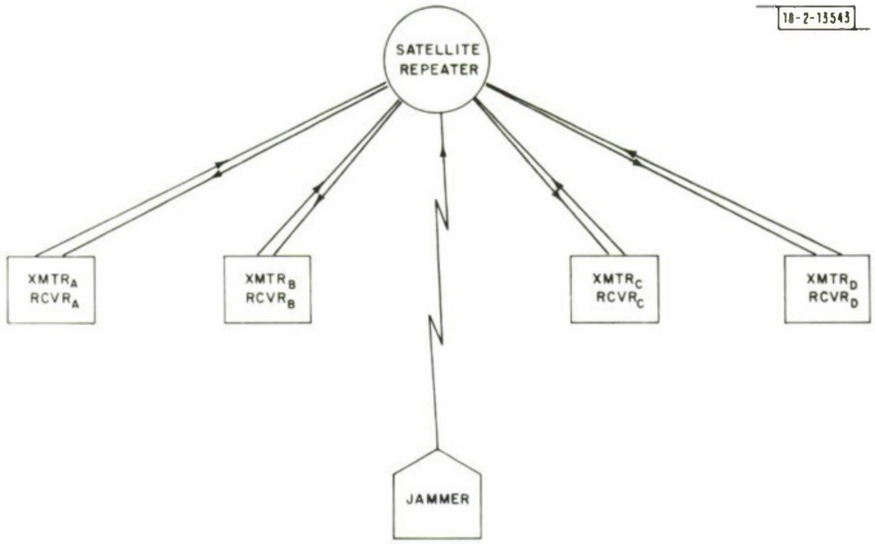

Fig. 2. Model of broadcast satellite channel.

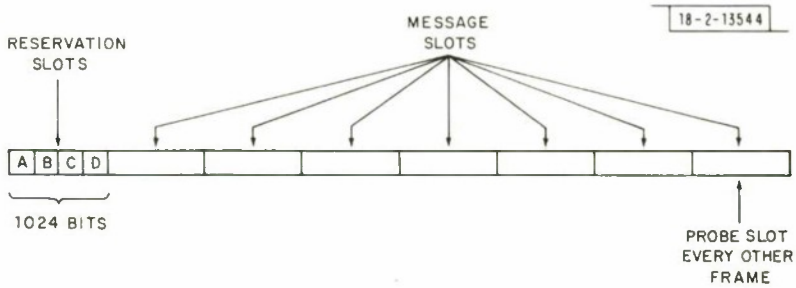

Fig. 3. Frame structure of broadcast satellite channel.

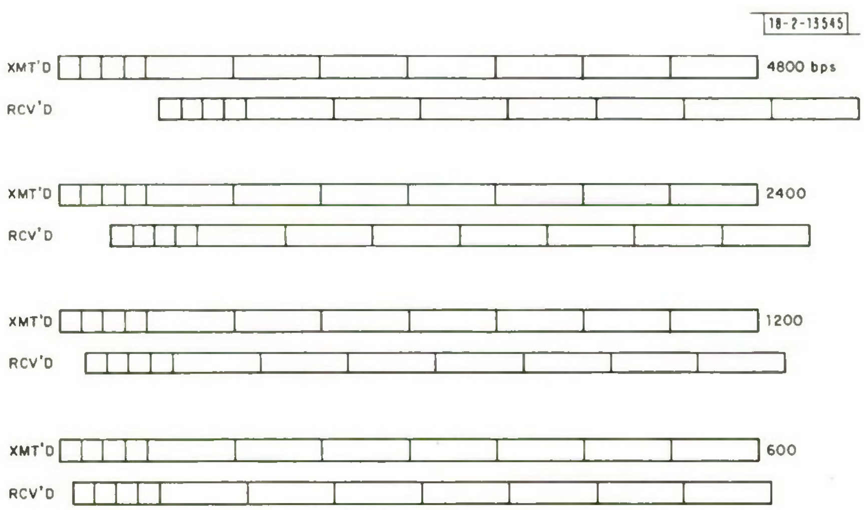

Fig. 4. Effect of round trip propagation delay on frame-by-frame basis relative to the bit rate of a satellite channel.

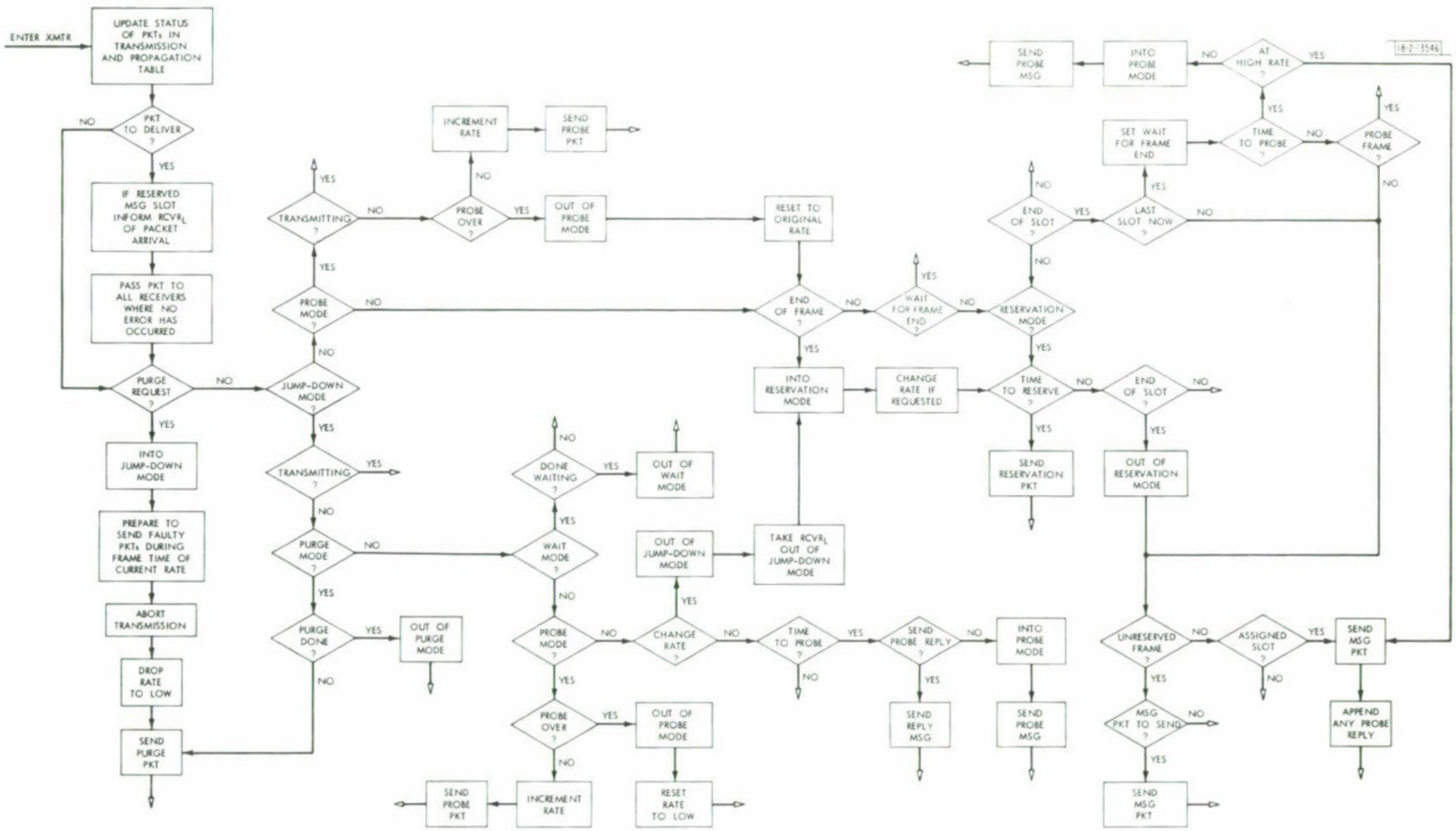

Fig. 5. Broadcast satellite channel transmitter protocol.

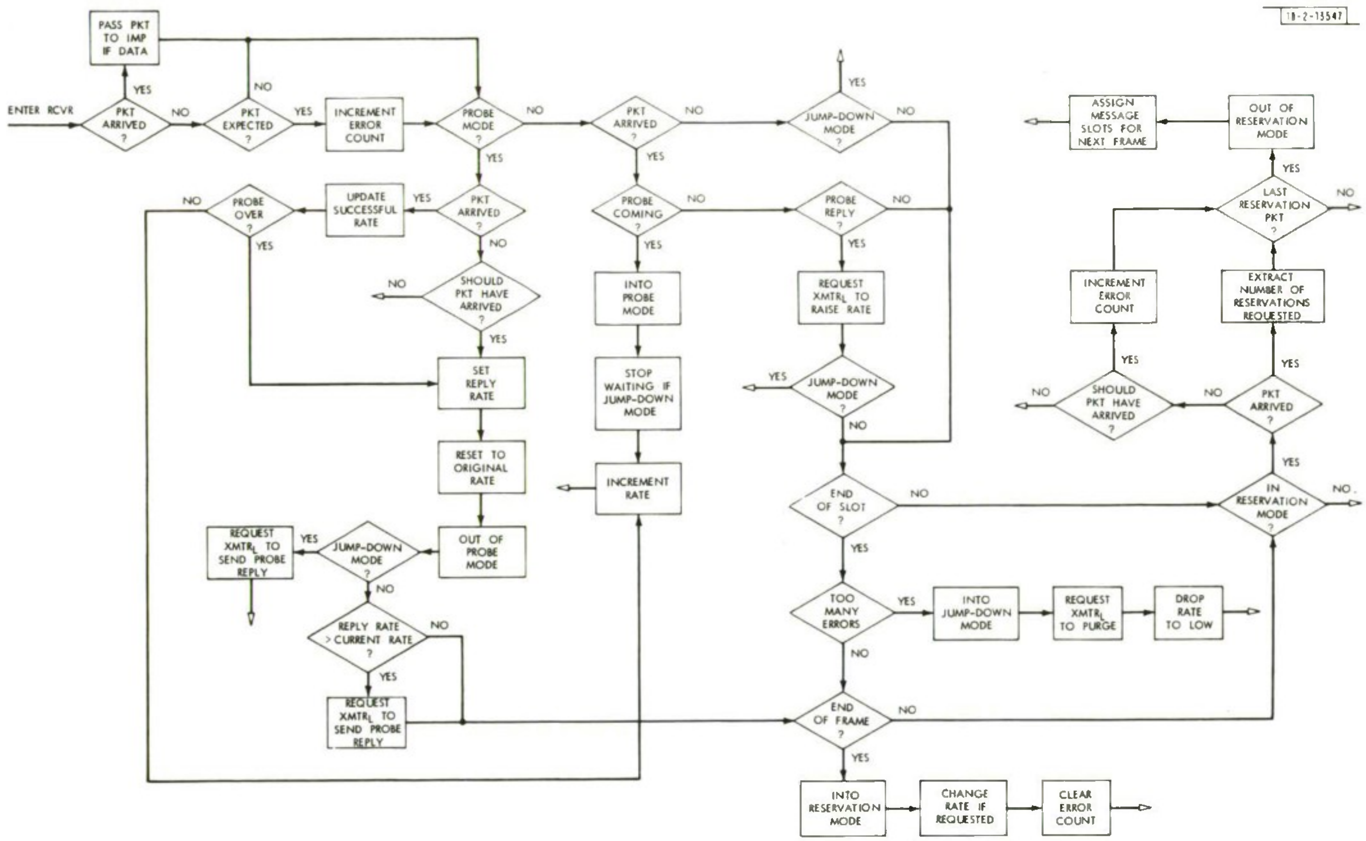

Fig. 6. Broadcast satellite channel receiver protocol.

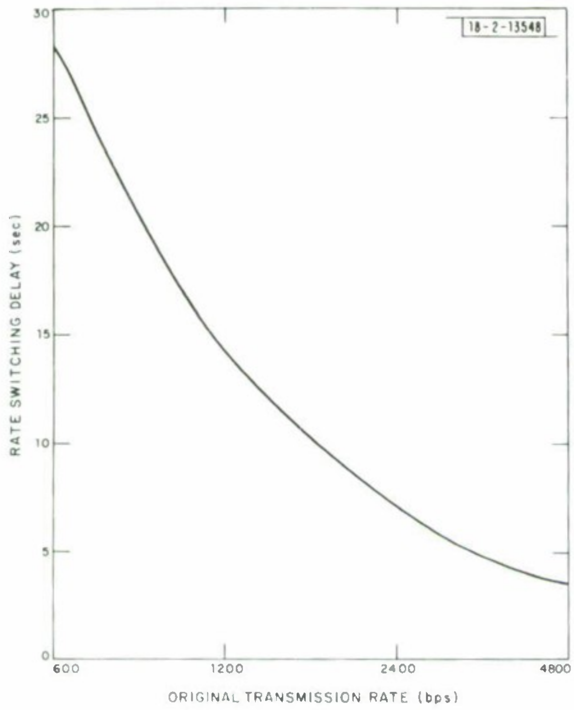

Fig. 7. Average-rate switching delay to capacity increase relative to original transmission rate of broadcast satellite channel.

Figure 7 shows average-rate switching delays after a capacity increase.

Adaptation to a decrease in channel capacity is accomplished by sensing errors. An error is defined as the absence of a correct reservation packet or the absence of a message packet sent by the local transmitter. If three errors are sensed during a frame, the receiver and its transmitter drop immediately to the lowest rate and deliberately incorrect packets are transmitted over a period of time which is equal to one frame at the original rate. This "purges" the system, insuring that all receivers will take similar action. This purge frame is followed by a "waiting" frame to allow forced-down transmitters to transmit a late purge frame. When the waiting frame is over, the transmitter monitors an external clock, waiting for its dedicated jump-down probe slot. Upon receiving a probe, all waiting transceivers will curtail their waiting frame and begin monitoring the external clock. A probe reply will be transmitted during a dedicated probe slot. Upon receiving a reply, all transceivers will resume normal framing at the new channel capacity. Downward switching delay is given by

$$T_{DD} = \underbrace{\frac{n}{2\mu_F C}}_{\substack{\text{average time}\\ \text{to detect errors}\\ \text{at receiver}}} + \frac{P}{2} + \underbrace{\frac{2n}{\mu_F C}}_{\substack{\text{time to}\\ \text{purge and}\\ \text{wait}}} + \underbrace{\frac{1}{2\mu_F C_{min}}}_{\substack{\text{average time}\\ \text{until first}\\ \text{probe}}} + \underbrace{\frac{K}{\mu_F C_{min}}}_{\substack{\text{time before}\\ \text{replying to}\\ \text{probe}}} + \underbrace{\frac{1}{\mu_C C_{min}}}_{\substack{\text{time to}\\ \text{receive}\\ \text{reply}}} + P$$

where

$1/\mu_F$ = full packet size in bits

$1/\mu_C$ = control packet size in bits

C = original channel capacity (bps)

C_{min} = minimum channel capacity (bps)

n = number of full packets per frame

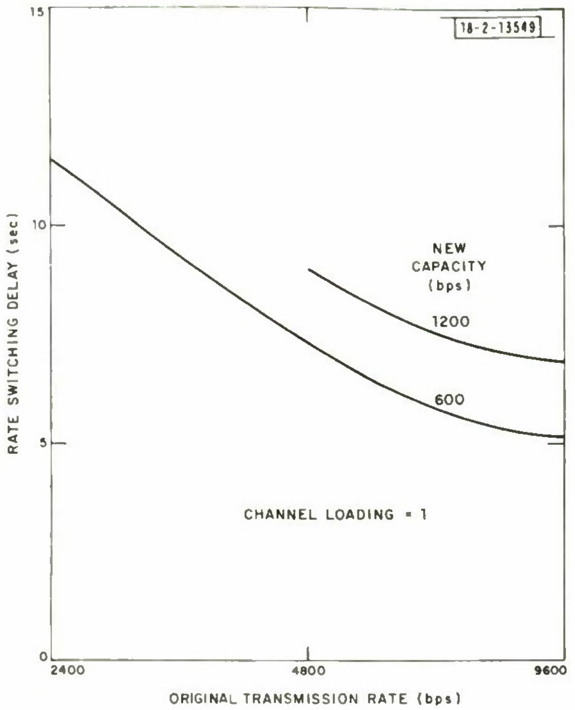

Fig. 8. Average-rate switching delay to capacity decrease relative to original transmission rate and new capacity of broadcast satellite channel.

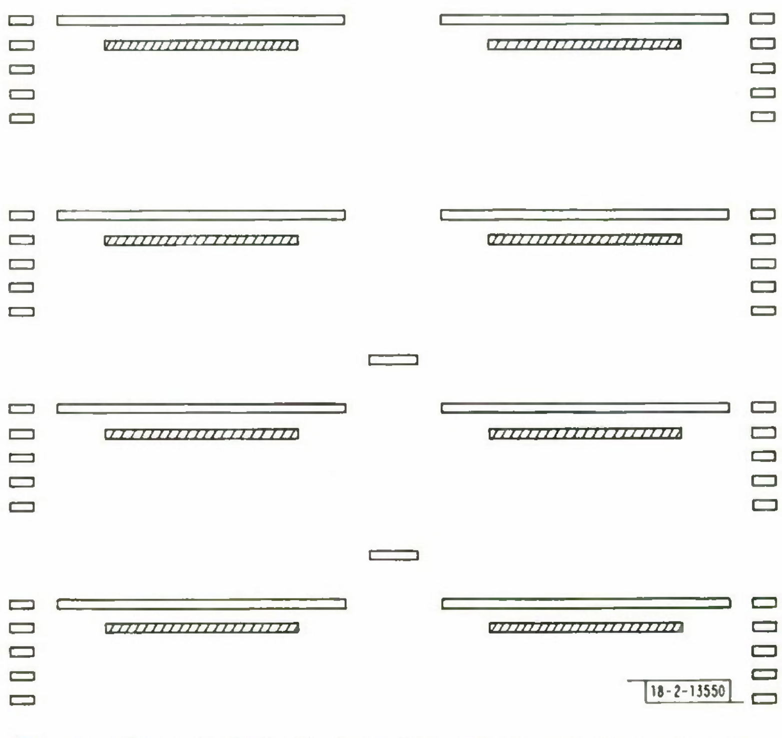

Fig. 9. Dynamic link display of broadcast satellite channel.

P = round trip propagation delay

$$K = \begin{cases} 1 & \text{if new capacity} = C_{min} \\ 2 & \text{if new capacity} > C_{min} \end{cases} .$$

Referring to Fig. 8, the difference in downward switching delay relative to channel capacity is explained by the fact that only at the lowest capacity will a receiver reply be available in time for the next probe slot.

The dynamics of the broadcast satellite channel model can be displayed on the Comtal color scope as shown in Fig. 9. The tick marks indicate the five bit rates, 9600 bps at the top and 600 bps at the bottom. Channel capacity is indicated by the long white bars, and transmitter and receiver rates by the shorter bars. All transmitters are shown at the left, opposite their local receivers at the right. The channel, as depicted, is operating at 4800 bps and is experiencing a collision of packets transmitted by transmitters B and C during an unreserved frame. If capacity were to drop, the observer would see broken lines appear, indicating errors sensed by the receivers.

III. DEMONSTRATION SYSTEM FOR PACKETIZED VOICE OVER A SINGLE-USER LINK WITH TIME-VARYING CAPACITY

A. System Description

1. Background

A demonstration system in which the packetized output of one variable-rate speech encoder is transmitted through a simulated radio channel with time-varying capacity has been developed and currently is being implemented on the PDP-11/45 and several peripheral processors. The demonstration facility will operate in real time and will allow a person to speak through the system while the capacity of the channel changes due to the addition of varying amounts of Gaussian noise to the channel signals (simulating jamming, fading, etc.). The system responds to capacity changes by appropriately changing the modem data rate and speech encoding bit rate. The subjective effects of the resulting disruptions and delays in speech can be determined for different degrees of channel degradation and various adaptation strategies.

This demonstration system provides a practical application for the variable-data-rate estimation procedures developed over the past two years (see Sec. II). In the present system, the protocols will be interfaced with a real-time simulation of a recently developed variable-rate modem (Sec. IV).

The demonstration system can be viewed as a general-purpose tool in which the subjective effects of adapting speech data rate to a varying channel capacity can be determined. The variation of channel capacity could result not only from channel noise due to jamming, but from a variety of mechanisms. For example, the delays and packet loss for one user's speech through a simulated multi-user network with varying voice and/or data loads can be measured from a non-real-time simulation. These measurements can then be applied to the demonstration system to determine the resultant subjective effects.

2. The Demonstration System

The half-duplex real-time variable-rate speech system to be implemented consists of several components. Referring to Fig. 10, a vocoder algorithm will reside in the LDSP which will

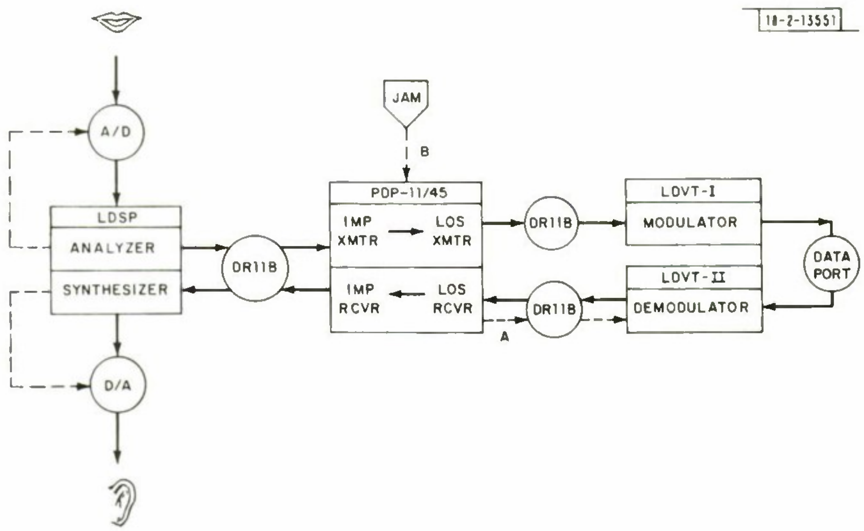

Fig. 10. Half-duplex variable-rate speech system.

analyze digitized speech input and send encoded speech data to the PDP-11/45 in a serial word stream via the DR11B, a direct memory access data transfer interface. The average rate of data transfer will be dictated by the A/D clock controlled by the LDSP.

The transmitting process modeled in the PDP-11/45 will form packets of 60 words of speech data and 4 words of header information consisting of packet delimiter, control information, and packet sumcheck. Packets will then be expanded to 128 words, each word containing 8 user bits, current transmission rate, and the noise level of the channel which will be an externally controlled parameter (labeled JAM in Fig. 10). These packets will be transmitted in a serial word stream via a second DR11B to a simulated modulator in LDVT-I. The average rate of this data transfer will be governed by a clock in LDVT-1.

The modulator will convert each group of 4 user bits (corresponding to 1 baud) into digital sinusoids of one of 16 frequencies, whose duration (termed the baud time T_b) depends on the current rate of transmission. This digital signal will be distorted as designated by the current noise level and sent directly to LDVT-II via a digital interface. The demodulator in LDVT-II will convert the received signal into bits and determine an upward rate assessment for this baud (see Sec. IV). The demodulator will then pass a serial word stream to the PDP-11/45 via a third DR11B, each word consisting of 8 data bits and two upward rate assessments. The dotted line marked A in Fig. 10 represents a periodic flow of control data from the PDP-11/45 to the LDVT-II which consists of the current rate of receiving. The receiving processor in the PDP-11/45 will then assemble the received data bits into packets, strip off the 4-word headers, and send a serial word stream to the synthesizer in the LDSP which will generate digital speech for the D/A converter. The receiving processor will update its channel capacity estimate based on the upward rate assessments by the demodulator. Capacity will be reassessed every 1024 bits; and when the receiver determines an increase in capacity, it will "send" the proposed rate to the transmitter. In the half-duplex system, the time to send a message over the imaginary return link is assumed to be negligible if the link is operational. When the transmitter receives the proposed rate change, it will append a rate hike message to the next packet

it transmits and change its rate after that packet has been transmitted. The receiver, upon receiving the rate hike message, will change its rate, a higher-rate vocoder will be loaded into the LDSP, and communication will continue.

The receiving processor also will monitor sumcheck errors and missing packets. Speech data will be passed to the synthesizer, regardless of sumcheck errors, given that an error does not occur in the packet delimiter. When the noise level becomes intolerable, the receiver will drop to the lowest rate and "send" the transmitter a request to drop its rate. When channel communication resumes at the lowest rate, channel capacity will be established, a new vocoder will be loaded into the LDSP, and communication will resume at the new rate. Table I shows

TABLE I

MODEM AND CORRESPONDING VOCODER BIT RATES

Modem Rates (bps)	Vocoder Rates (bps)
20,480*	19,200 APC
10,240*	9,600 APC
6,826	6,399
5,120*	4,800 LPC
4,096	3,840
3,413*	3,200
2,925	2,742
2,560*	2,400 LPC

* Rates for channel operation.

the range of rates which the modem will support. Rates were derived by letting the baud time take on values $T_b = kT_{b_{min}}$ where $T_{b_{min}}$ is the time necessary to transmit 4 bits at 20,480 bps. The modem will make upward rate assessments to 40,960 bps. Because an estimate of channel capacity over only 1024 bits ensures "error-free" (probability of error $< 5 \times 10^{-5}$) transmission at a rate equal to one-half capacity, the channel will be operating at one of the five rates marked by asterisks. These channel rates will accommodate the vocoder rates indicated. 2400- and 4800-bps linear predictive coding (LPC) algorithms will be supported as well as 8000- and 16,000-bps adaptive predictive coding (APC) algorithms padded to 9600 and 19,200 bps.

The mechanisms for forming packets, expanding packets for the modem, reassembling packets, monitoring errors, and adapting to channel capacity have been implemented and checked out on the PDP-11. Flow charts depicting the operation of transmitter and receiver are shown in Figs. 11 and 12.

The PDP-11 control program that can selectively bootstrap one of several prestored vocoder algorithms into the LDSP has been written. The program has been debugged on an LDVT and the algorithm switching transient is almost imperceptible.

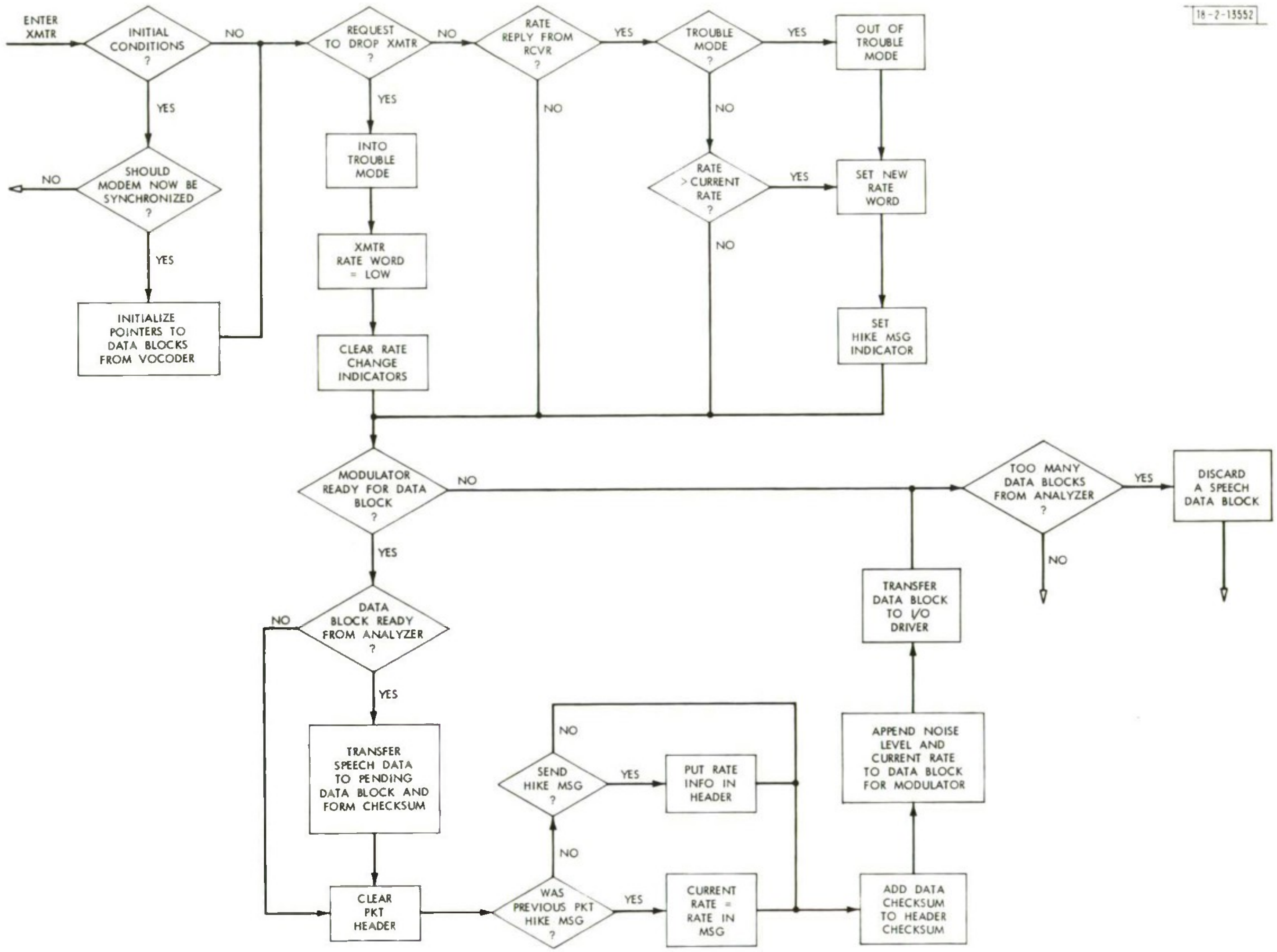

Fig. 11. Transmitter for half-duplex variable-rate speech system.

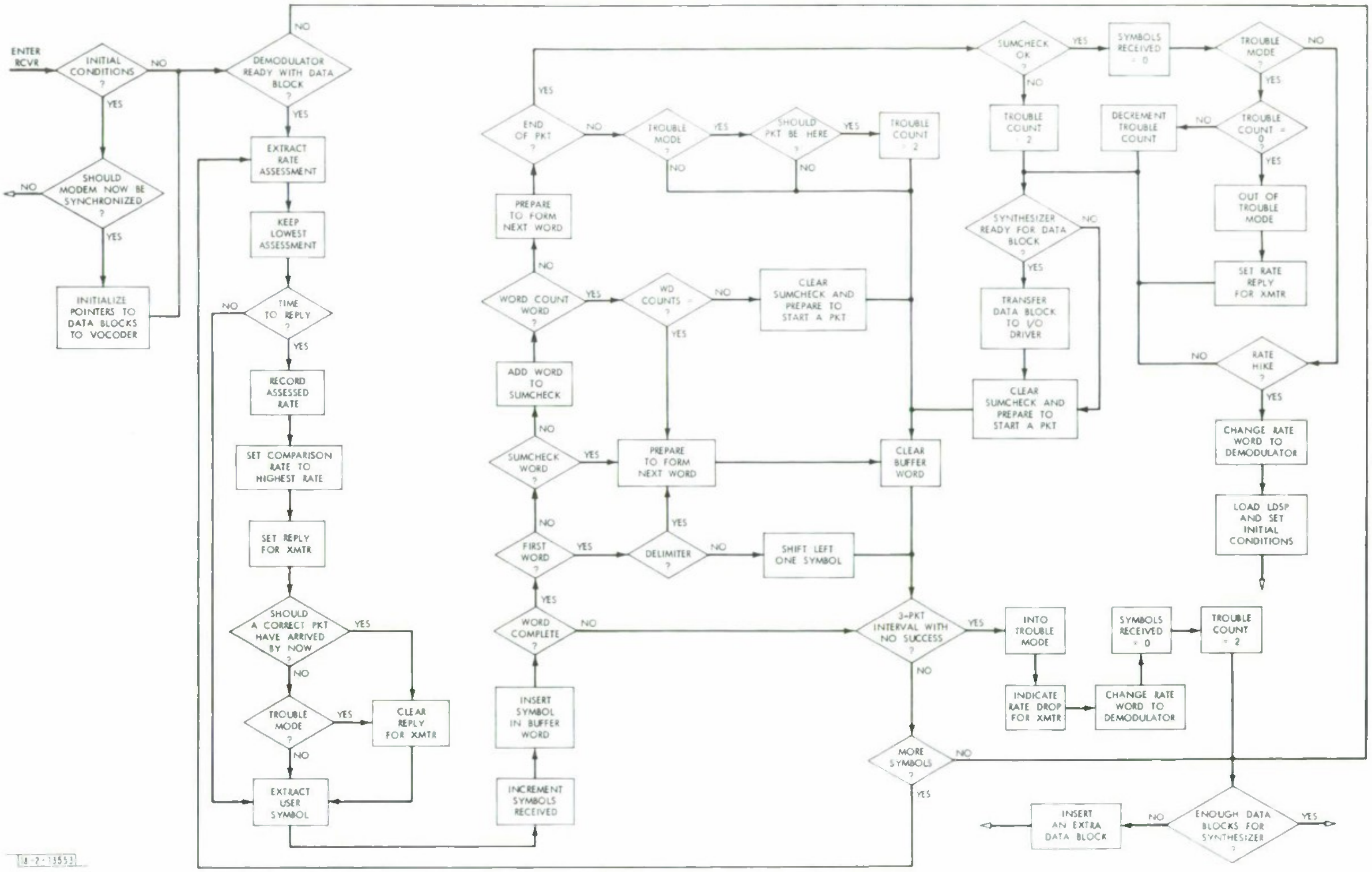

Fig. 12. Receiver for half-duplex variable-rate speech system.

B. Flexible Signal Conditioner

The speech encoder simulated in the LDSP changes vocoding rates by switching vocoder algorithms. A library of LDSP vocoder algorithms is stored in the PDP-11/45: any one can be selectively bootstrapped into the LDSP under PDP-11 control. Since not all algorithms employ the same audio pre- and post-filtering or A/D clock rates, a flexible, computer-controlled signal conditioner has been designed and is presently being constructed. The design of the signal conditioner allows the independent selection of a variety of low-pass pre- and post-sampling filters, pre- and de-emphasis filters, A/D and D/A clock rates, and audio signal attenuation. The device will be controllable from one of several sources: front panel switches, the LDSP, the PDP-11, and two LDVTs.

A block diagram of the signal conditioner is shown in Fig. 13. There are separate data and control channels between the controlling processor and the device. The processor can read such status parameters as A/D overflow. The signal conditioner includes all the audio circuitry and digital interfacing required to support the LDSP's role as a general-purpose real-time speech-processing facility. It also will provide additional audio I/O flexibility to the LDVTs and the PDP-11/45.

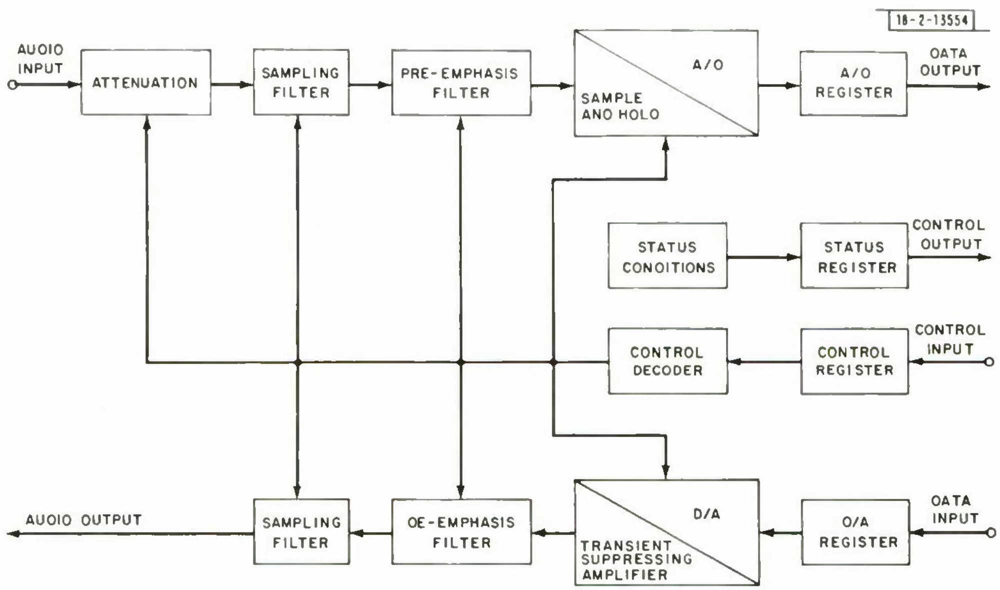

Fig. 13. Flexible signal conditioner.

IV. ADAPTIVE VARIABLE-RATE MODEM WITH "NO-COST" UP-PROBING

A. Background

Previous work in the Airborne Command and Control program led to the development of variable-rate protocols for use with line-of-sight and broadcast satellite channels. A basic assumption in these protocol designs was that the network processors in which the protocols resided had access only to the demodulated or decoded streams of user data, and were essentially ignorant of the detailed nature of the link degradation mechanisms. This resulted in strategies that required the explicit transmission of packets at higher than the current channel bit rate in order to assess whether the links could support increased data loads. The cost of

these "up-probing" transmissions was on the order of ten percent of current channel capacity, thereby resulting in a commensurately reduced user-available bandwidth.

In the process of designing the components of a real-time implementation of sustained packet voice communication over a simulated time-variable line-of-sight link, it became evident that, with judicious modem designs, one could eliminate the need for explicit up-probing for channel assessment purposes. As with the original protocols, link degradations would continue to be determined at the nodal processor level; however, the modem receiver could provide the upward assessment function via observation of the incoming channel signals. This leads to a more efficient use of the channel and, additionally, to a simpler and more dynamically responsive protocol.

B. Up-Probing Technique

We assume that channel signals or "baud" are sine waves of known frequency and/or phase, and of duration T_b seconds. We further assume that when the data transmission rate changes, only the value of the baud time T_b is affected, i.e., the phases or frequencies that are used are identical for all modem rates. In receiving these signals, the modem receiver performs a number of correlation operations, in each case integrating over the duration, T_b, of a channel signal. The number of matched filters required is simply the number of distinct orthogonal signaling waveforms that have been defined, e.g., M for an MFSK system, two for a QPSK modem, etc. The basic observation that allows one to effect a no-cost up-probing function is that in the process of integrating over T_b seconds, and assuming they are "dumped" after each such interval, the filters go through a continuum of states that are identical to what they would have produced had they been directed to integrate for less than T_b seconds. Since it is assumed that the current rate is essentially error-free (otherwise it would have been changed by the protocols), one way of probing the channel at, say, twice its rate would be to "remember" the outputs of the matched filters at $T_b/2$ and compare them to their values at T_b. If the detection logic reaches the same conclusion (regarding frequency or phase, etc.) for both cases, one can conclude that this particular baud could have been successfully communicated at twice its current rate. Obviously one can probe up to any rate in excess of the current one and in fact the extension to multiple-rate up-probing is trivial. The important element here is that the key parameters of the signaling waveform remain the same for all rates and only the baud durations be allowed to vary. This notion supports a variety of signaling strategies such as FSK, MFSK, QPSK, DCPSK, etc., and our particular choice of MFSK has been made somewhat arbitrarily and in anticipation of computational efficiency through the use of the FFT.

Although the modem can assess the highest rate at which each baud could have been successfully received, it cannot without additional information declare that the channel is in fact usable at that rate. That decision depends on a number of considerations (such as packet error rates) that are best left to the nodal processor. Our strategy, therefore, is merely to deliver to the node a baud-by-baud upward assessment and to allow the ultimate rate changing decision to be made at the nodal processor level.

C. Effect of Antijam Signal Modification

In some applications a baseband signaling waveform is further modified before transmission in order to provide jam-resistance. This is often accomplished by subdividing the baud into smaller time units and sending each unit at a different frequency (frequency hopping) or with

different phase. The hopping or phase-shifting sequences are generally of pseudorandom nature, thereby spreading the spectral content of the transmitted signal uniformly over a given band. The particular pseudonoise sequence is known both to the transmitter and receiver, and at the receiving end the wideband input is appropriately unraveled to yield the original baseband signals. These can then be filtered and detected in the normal way. One can therefore think of the spread-spectrum process as one that can be transparent at the baseband level, and the above-described up-probing technique, since it deals with baseband signaling waveforms, can still be used in an A/J environment.

As a practical matter, it may not always be true that phase coherence of the baseband signal is maintained in the spread-spectrum process. This is especially so in frequency-hopping systems, for which propagation characteristics may be different in the various frequency bins. The major effect that this has on the baseband matched filters is that they may be required to combine both coherent and noncoherent integration methods over a T_b-second interval, rather than to perform the exclusively coherent summation they would in the absence of hopping. If one constrains the A/J device to hop or switch at the same rate (e.g., constant chip rate) regardless of the value of T_b, it again turns out that the upward assessment function can be performed as described. The difference, of course, is in the mix of coherent and noncoherent processing that is used in the baseband matched filters, but this is required independently of the probing activity.

An assumption in the above argument has been that the A/J portion of the modem receiver is of the correlation type, playing stored replicas of the pseudonoise sequence against the incoming data stream at the appropriate times. This converts the input to its original baseband form prior to final filtering. In some cases a more attractive receiver consists of single filters that are matched to the entire incoming waveform. The ARPA packet radio network, for example, will use surface-acoustic-wave filters that match the aggregate signal consisting of a differentially coherent PSK baseband waveform modified by a pseudorandom phase alternation sequence. In these filters the outputs consist of narrow pulses occurring at the baud rate, and the notion of prematurely sampling the filter outputs for upward assessment purposes is void. The basic information required for up-probing is still contained in the channel signals, however, and one might consider using a number of independent SAW devices, each matched to successively shorter segments of the longest expected baud. These would be necessary to support multi-rate operation anyway, and given that they are included in the receiver, they could be exploited for upward assessment purposes in much the same way we have proposed. Although it is not clear that variable-rate issues are currently of particular concern in the packet radio network, or that the eventual economies of SAW devices make multiple front ends a viable option, the exercise is offered as an example of how our up-probing technique can be employed in a more realistic and complicated situation than exists in our laboratory experiment.

D. Modem Implementation on the LDVT

The transmit and receive modems will be implemented in software on two separate LDVTs. The transmit modem is fairly trivial to implement in that it consists of a simple table lookup routine to produce the samples of the transmitted sinusoid. The receive modem is much more complicated to implement because it requires demodulation, up-probing, and synchronization algorithms.

The basic demodulation algorithm is a 16-point FFT. At the lowest transmission rate ($\approx$2.5 kbps) 16 such 16-point FFTs are taken on adjacent data points, their outputs are coherently added, and the magnitudes of the resulting data are compared to determine the largest.

The demodulation strategy just described can be made much more efficient computationally at lower rates by properly adding FFT inputs rather than outputs. At the highest rate, pairs of data samples spaced by 16 samples are added and then the FFT is taken. Similarly, at the lowest rate, 16 groups of 16 data samples separated by 16 samples are added and the 16-point FFT performed on the resulting sums. This procedure results in the same demodulation algorithm described above but requires many fewer FFTs, thus resulting in an enormous computational saving.

The computational savings just described cannot be fully realized because of the modem's requirements for up-probing and synchronization. Up-probing requires that demodulation be done on all possible baud lengths smaller than the current baud lengths. This could be accomplished by taking individual 16-point FFTs and adding them in various combinations to obtain the demodulated output for the smaller baud lengths as well as the demodulated output for the current baud length. This procedure throws away the computational savings that can be gained by adding the data before taking the FFT. Some of this saving can be regained by the following strategy. The data points comprising the first half of the baud are added in the manner described above and an FFT is taken. Next, one half of the points in the remaining half of the baud are added as above and another FFT is taken. This procedure is continued until all the data comprising the baud have been used up. If the baud consists of 16M points, this procedure results in a total of $(1 + \log_2 M)$ FFTs being taken as opposed to the M FFTs required when no pre-adding is done.

The above procedure yields all the data needed to demodulate bauds smaller than the current baud with what appears to be a minimum of computational effort. In addition, all information needed to perform synchronization is also present. The synchronization algorithm consists of computing 16-point FFTs for each half of the current baud and deriving a difference signal from the FFT output points corresponding to the whole-baud FFT output point that had maximum magnitude. This difference signal is smoothed and used to shift the baud boundary forward or backward in time.

V. ADAPTIVE VOICE COMMUNICATIONS ON A TIME-VARYING, INTEGRATED NETWORK

A. Introduction

The demonstration system described in Sec. III provides a vehicle for studying the subjective effects of adaptive-rate speech communication with a time-varying channel capacity between voice terminals. A more long-range objective is to develop adaptive speech communication techniques in the context of a large, integrated network including both satellite and terrestrial links. The links in such a network would generally be wideband and would be shared by many users. Therefore, the effect of a change in link capacity would be distributed among many users in a manner determined by the network nodes. In addition, the channel capacity available for any particular conversation would be affected by the fluctuations in competing traffic, even if all communication links in the network retained constant capacity.

We describe below a networking strategy for handling variable-rate packet speech communication in the context of a large, integrated network. This strategy represents a three-way

marriage between a Lincoln-developed PVC concept,[1] an embedded speech coding technique of the type under investigation at the Naval Research Laboratories, and variable-rate communications protocol similar to those described in Sec. II above. A new dual-rate APC speech algorithm which is very well matched to the proposed networking strategy is described. Finally, a simple model for the traffic generated by many voice users sharing a single wideband link is defined. From this model, an expression is derived for the fraction of packets lost, representing the channel degradation seen by a typical voice user due to competition with the other users.

B. An Adaptive Variable-Rate Packet-Speech Strategy

1. Overview

A rate-adaptive packet-speech network design based on a PVC network concept, an embedded speech coding technique, and a variable-rate communications protocol, is described. Potential features of the design are:

(a) Simplicity of nodal switching logic,

(b) Continual adaptation of voice user bit rates to time-varying network traffic conditions,

(c) Ability to accommodate users of different individual priority levels such that when bit rate reductions are required they will tend to be directed at the lower priority speakers,

(d) Minimal network supervisory activity,

(e) Simultaneous accommodation of a variety of speech algorithms and user adaptation capabilities within the network.

In brief, voice subscribers view the network as a time-variable communications medium and attempt, via a simple user-to-user protocol, to maintain communications at the highest rate that the network will currently allow. Processing at the network nodes is minimal, with each node selectively discarding voice packets on the basis of packet priority levels and the instantaneous conditions prevailing in its voice and data packet queues.

2. Network Speech Terminal

Network speech terminals would produce priority-ordered packets that are compatible with an embedded coding form of speech synthesis. That is to say, a high rate (16 kbps) vocoder frame consists of a number of distinct subsets of bits such that one of them can be used for synthesizing low-rate speech (say 2.4 kbps); that same subset when used in conjunction with one of the others supports an intermediate rate synthesizer (like 4 kbps); the first two in combination with a third yield 8-kbps results, etc. Each distinct subset is transmitted in a separate packet whose assigned priority is a monotonically decreasing function of the lowest synthesis rate for which its contents are used, i.e., higher priorities are accorded to packets that support lower-rate synthesis. Whether or not a given terminal actually transmits all its generated packets will depend on its knowledge of the data rate that the network is currently supporting. Thus, if it is known that speech rates in excess of 4 kbps are unlikely to be received successfully, the terminal need only generate packets of its two highest priority types. An end-to-end protocol that provides this information is described below.

In its receiving mode the network speech terminal receives priority-ordered packet types from the network and effects a frame-by-frame synthesis at whatever rate the received sequence of packets will support for that frame. Thus, if in one frame a full set of packet priorities is received, the terminal produces wideband synthetic speech. If in another only the two highest priority levels are received, a 4-kbps synthesizer is invoked. Receipt of only the highest priority packet will result in 2.4-kbps synthesis, and so on. The action of the network is such that upon sensing impending overload conditions, the nodes discard speech packets of lower priority first, thereby tending to reduce the bit rates of wideband users while attempting to maintain a continual flow of at least narrowband data to all users.

3. Nodal Strategy

The handling of voice packets by the network nodes is based on the PVC concepts. We assume that dual queues (voice, data) exist at each node for every trunk line leaving that node. Data packets are transmitted when conditions in the voice and data queues satisfy appropriate criteria, e.g., number of packets in each queue, etc. With respect to the voice queue, we imagine that at a given instant of time the packets awaiting service are of a variety of priority types. Depending on the length of the voice queue and perhaps additionally as a function of the length of the data queue, service may be denied to voice packets of lower than a given priority level. Denial of service results in loss of the packet. The priority cutoff thresholds vary with time based on the observed lengths of the voice and data queues, and are governed by local algorithms that function independently in the various nodes. When queues are of moderate length, no packets are discarded.

Packet elimination may be accomplished by any of a number of mechanisms. One can admit all packets into the queue and eliminate those of low priority at the time they would normally require service, i.e., at the queue output. This basically reduces the service time for these packets to zero, thereby speeding up the flow. Alternatively, one can deny entry into the queue for low-priority arrivals, resulting in a discouraged arrival type of strategy. The first method has the advantage of not discarding a packet unless it is absolutely necessary, but it results in longer queues when large numbers of packets are being discarded. A combination of the two, in which one queues packets of higher than one or two priority levels below what is currently being discarded at the output, might perform better than either method alone. A third mechanism is one in which the contents of the queue are examined and packets removed from its interior. This might be valuable at times of impending catastrophe.

4. Adaptation Protocol

Two speakers engaged in a point-to-point conversation see the network as a mechanism that introduces sporadic packet loss that tends to lower the bit rates at which they communicate. Because of the priority-ordering of their speech packets, however, they rarely lose communication. Instead, the bit rates (and thus the fidelity) of their transmissions vary with time. Obviously, each speaker can transmit at the widest bandwidth of his terminal and allow the network to reduce his rate to a more realistic value. This strategy, however, places unnecessary burdens on those nodes preceding the one (or more) at which heavy traffic conditions prevail. Since in a PVC network routing is established at the time of initial connection and is based on the traffic loads throughout the system, a better match between transmitted and received bit rates would be desirable. This can be accomplished via a very simple end-to-end protocol. In fact, all that

is needed is that the listener periodically inform the speaker of the average rate at which he is currently receiving transmissions.

Our strategy, therefore, is to have each speech synthesizer (or its associated host) "remember" the average rate at which it has most recently been producing speech, and periodically inform its counterpart at the other end of the connection of this value. The remote analyzer in turn responds by appropriately discarding packets before they enter the network, thereby matching its rate to the value that is currently being sustained for that particular connection. Since point-to-point conversations are two-way (but not necessarily identically routed), the observed rate declaration from a given terminal can conveniently be appended to one of its own transmitted speech packets. One would probably use the lowest bit rate (highest priority) voice packet type for this function, resulting in an additional overhead of on the order of only several bits per second.

Depending on the specifics of the embedded coding algorithms, annoying perceptual effects may result from frequent switching of synthesizer rates. Thus, if an isolated packet is lost in the network causing, say, a single frame in a normal 8-kbps stream to result in 4-kbps synthesis, the receiver may prefer to remain at the lower rate for the remainder of the current talkspurt or for a specified period of time. These notions can be embodied in the speech terminals without modification of the networking strategy. If as a result of sporadic packet loss, however, a synthesizer remains at the lower rate for long periods of time, it can account for this in reporting to the other station, and cause the latter to drop to that lower rate.

There will be times when, having established communication at some low bit rate, network conditions relax, and it becomes possible for a pair of terminals to increase their rates. This can be automated via a probing strategy in which, when a transmitter observes that its data are being received at the same rate they are being sent, it tries sending faster. If the synthesizer fails to receive at the new rate, the transmitter, by virtue of the above-described rate-lowering strategy, drops back. Thus, all voice terminals probe the network by occasionally trying to send voice packets of lower priority than are currently being propagated successfully.

5. User Priorities

In order to provide varying grades of service to a variety of users, packet priorities can be modified on the basis of individual user identities. For example, if user A's speech terminal generates a vocoder frame of, say, four distinct data subsets resulting in four packet priority assignments, his packets might enter the network with priority values A_1, A_2, A_3, and A_4. User B, whose personal priority is lower than that of user A might, if he also produced a four-level embedded data stream, get levels B_1, B_2, B_3 and B_4 such that the overall ranking of the eight packet types was A_1, A_2, B_1, A_3, B_2, A_4, B_3, B_4, in descending order of absolute priority. If a network node, in coping with traffic overloads, was forced to discard all packets of lower than A_4 priority, then user A would get 16-kbps service while user B would operate at 4 kbps. Similarly, if a node had to discard all packets of lower than B_1 priority, user A would drop to 8 kbps and user B to 2.4 kbps.

As a final topic, we observe that the specifics of the speech algorithms, the number (if any) of levels of embedded coding, etc., are a private matter between two end-to-end users. The network deals only with packet priority values, which are assigned either in advance for each user or by the network when a connection is established. The network can thus support a variety

of speech algorithms and embedded coding levels without special accommodation. In the case of fixed-rate users for which no embedded coding exists, a judicious choice of packet priority will either guarantee communication at the fixed rate, or result in occasional packet loss (as in the original PVC concept) when traffic volume is high.

C. Dual-Rate Adaptive Predictive Coder

A dual-rate APC algorithm, suitable for embedded coding and speech transmission via priority-oriented packets, has been developed. The scheme involves a standard APC vocoder with the option of transmitting the error signal at two bits per sample for higher fidelity and at higher rate, or at the usual one bit per sample. A block diagram of the dual-rate APC transmitter is shown in Fig. 14. For network speech transmission, packets of higher priority

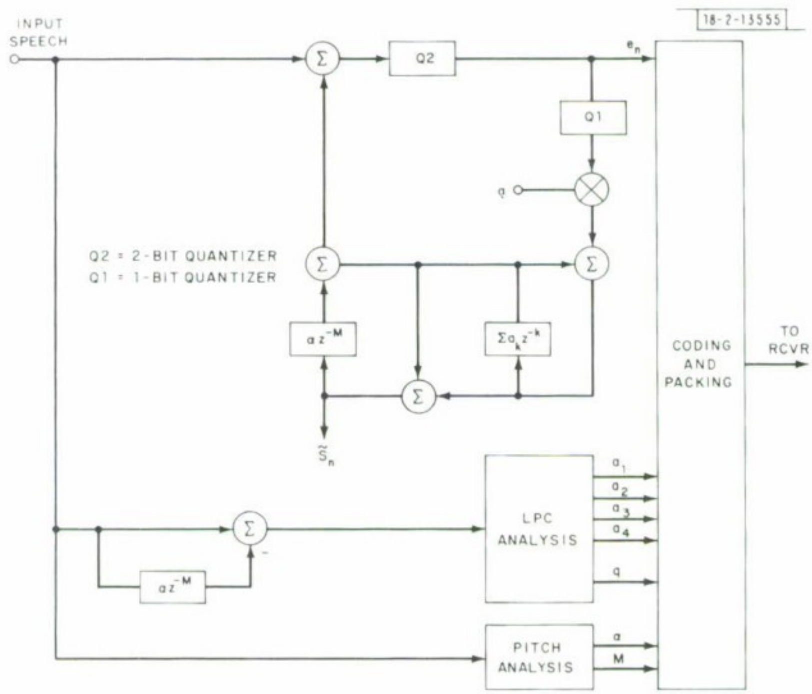

Fig 14. Dual-rate APC transmitter.

would contain the most significant bit of each 2-bit error signal sample, plus the LPC, gain, and pitch parameters. The lower priority packets would contain the less significant bits of the error signal samples.

An APC system based on this dual-rate technique has been implemented on the LDVT. The speech bit rate for one-bit error signals plus LPC, pitch, and gain is 8 kbps, and the second error signal bit yields a higher rate of about 15 kbps. The quality of the high rate system was judged to be very good, and certainly far superior to 16 kbps CVSD (continuously variable slope delta modulation). No annoying transient sounds seem to result from switching between the two rates.

Thus dual-rate APC system offers a capability for rapid switching of high-rate users (16 kbps) down to a medium rate when network overloads begin to occur. A promising approach

for extending this technique to lower rates is the LPC-APC approach. The analyzer would be extended to produce more LPC coefficients. Synthesis at rates of 4 kbps and below would not require an error signal, but would employ only the LPC coefficients and pitch.

D. Talker Activity Model and Fractional Packet Loss for Statistically Shared Wideband Voice Link

1. Background

In a large integrated network, the capacity available for any particular voice conversation will be affected by fluctuations in competing traffic. We present here a talker activity model useful for predicting the effects of such traffic in a system (such as a packetized voice system) which takes advantage of speech activity detection. The application of this model to network analysis depends on many factors such as network topology, voice/data mix, and priorities and rates of the various voice users. However, in the particular case of many voice users of identical priority sharing a single wideband link, a formula for percentage packet loss can be derived. As the number of users gets too large for the link capacity, the packet loss probability will increase and an adaptive procedure such as that outlined in Sec. B-4 would have to be invoked.

2. Talker Activity Model

Silent intervals typically represent more than half the elapsed time in conversational voice transmissions. By using speech activity detection and rapid circuit switching, the TASI system[2] allows multiplexing of M talkers onto $c < M$ circuits with little degradation for suitably chosen c/M. Similar advantage can be gained in a packetized voice network by transmitting packets from each talker only when activity is detected.

In order to analyze or simulate a system which takes advantage of speech activity detection, a model for talker activity is needed. A simulation model[3] previously used in analysis of a PVC network represents M independent speakers, each switching between talkspurt and silence according to measured duration distributions. A simpler model is presented here for the random process $n(t)$ representing the number of talkers active at a given time. The details of talkspurt and silence duration distributions need not be known, and the behavior of talkers need not be simulated on an individual basis. In addition, the model provides a direct handle on the behavior of the key parameter $n(t)$. On the other hand, the model involves an assumption of large M which is not necessary in the earlier formulation.

Assume that we have M talkers, each switching alternately between talkspurt and silence. We do not consider here the effect of conversations beginning and terminating. Let

$$1/\lambda = \text{mean silence duration}$$
$$1/\mu = \text{mean talkspurt duration}$$

for each talker. The behavior of $n(t)$ will be characterized by step changes of ± 1 at random times. If M is large, then the pooled point process representing transition events (both up and down) in $n(t)$ is approximately Poisson, and the inter-transition intervals are exponentially distributed. This reasoning leads to a birth-death model for $n(t)$ with $M + 1$ possible states, and a transition rate diagram as depicted in Fig. 15. From state k, there are k independent

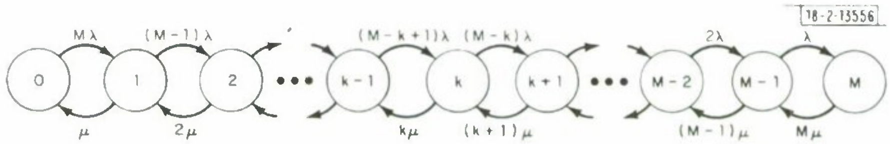

Fig. 15. State-transition-rate diagram for number of active talkers among M independent talkers each switching alternately between talkspurt and silence.

sources, each of average rate μ, for the event $k \rightarrow k-1$, representing one of the active talkers dropping into silence. Similarly, there are $M-k$ rate λ sources for a new talkspurt. Thus

$$\lambda_k = \lambda(M-k) \quad , \quad 0 \leqslant k < M$$

$$\mu_k = k\mu \quad , \quad k = 1, 2, \ldots, M \tag{1}$$

as indicated in the diagram. It turns out that this model for the number of active talkers is identical to the $M/M/\infty//M$ model in Kleinrock[4] for a finite customer population and unlimited number of servers.

The steady state probability that $n(t) = k$ is given as[4]

$$p_k = \frac{(\lambda/\mu)^k \binom{M}{k}}{(1+\lambda/\mu)^M} \quad . \tag{2}$$

The mean of $n(t)$ is

$$\overline{N} = \frac{M\lambda/\mu}{1+\lambda/\mu} \quad . \tag{3}$$

Assuming that the system is in state k at a given time, the time until the next transition will be exponentially distributed according to

$$f_{\tau_k}(\tau) = (\lambda_k + \mu_k) \exp[-(\lambda_k + \mu_k)\,\tau] \tag{4}$$

where

$$\lambda_k + \mu_k = (M-k)\,\lambda + k\mu = M\lambda + (\mu - \lambda)\,k \tag{5}$$

is the reciprocal of the mean time until the next transition. Note that in the special case $\lambda = \mu$ (mean talkspurt duration = mean silence duration), we have

$$f_{\tau_k}(\tau) = M\lambda\, e^{-M\lambda\tau} \tag{6}$$

independent of k. Again starting from state k, the probability that the next transition will be upward to $k+1$ is

$$P[k \rightarrow k+1] = \frac{\lambda k}{\lambda_k + \mu_k} = \frac{(M-k)\,\lambda}{(M-k)\,\lambda + k\mu} = 1 - P\,[k \rightarrow k-1] \quad . \tag{7}$$

A simple simulator for generating sample functions of $n(t)$ could be based on Eqs. (4) and (7). Starting from state k, the time until the next transition is drawn from an exponential distribution [Eq. (4)], and an independent up-down decision is determined according to the probabilities

given in Eq. (7). It might be of interest to carry out a mixed simulation, where the gross traffic pattern is modeled as just described and the talkspurt and silences of one particular talker are modeled in detail. Then a history of packet loss, delays, etc. for the individual talker could be collected, and subjective effects could be evaluated by imposing these losses and delays on real speech, using the facility described in Sec. III above.

It is interesting that the formula [Eq. (2)] for p_k can be obtained by much simpler considerations. Let p represent the probability that any given talker is active at a random time. Then since talkers are independent, the probability that exactly k of M talkers are active is the binomial

$$p_k = \binom{M}{k} p^k (1-p)^{M-k} \quad . \tag{8}$$

The result is identical to Eq. (2) with

$$p = \frac{1/\mu}{1/\lambda + 1/\mu} = \frac{\lambda/\mu}{1 + \lambda/\mu} \tag{9}$$

representing the ratio of mean talkspurt duration to the sum of mean talkspurt duration and mean silence duration. Equation (8) has been used successfully in estimating the number of circuits needed to handle M talkers in a TASI system.

3. Packet Loss on a Shared Wideband Link

Consider M talkers each switching between talkspurt and silence as in the model described above. Assume that these talkers each transmit packets at the same uniform rate (one every T_p seconds) while in talkspurt and that no packets are transmitted during silence. Let these talkers share, with equal priority, a single transmitting channel whose capacity is such that a maximum of c packets can be transmitted every T_p seconds.

Up to M packets can be produced every T_p seconds, and up to c packets can be handled by the channel. We will assume that the channel sharing scheme is such that if $k > c$ packets are produced in an interval of length T_p, $k - c$ of these packets, chosen at random from among the k active talkers, will be discarded. If $k < c$ packets are produced in an interval, all will be transmitted. We choose not to consider retaining the overflow packets in one interval for transmission in the next interval because this (a) significantly complicates the analysis, and (b) would add to the channel utilization only when k was fluctuating very closely around c, and probably provide little long-run advantage in utilization.

Given the above assumptions, the fraction of packets lost can be calculated quite easily. Over a long time of duration NT_p, the total number of speech packets offered by the M talkers will be

$$\text{packets offered} = \sum_{k=0}^{M} kNp_k \tag{10}$$

where p_k is given in Eq. (8) above. The number of packets discarded will be

$$\text{packets discarded} = \sum_{k=c+1}^{M} (k-c)\, Np_k \quad , \tag{11}$$

and the resulting fraction of packets lost is

$$\varphi = \sum_{k=c+1}^{M} (k-c)\, p_k \Big/ \sum_{k=0}^{M} k p_k$$

$$= \frac{1}{Mp} \sum_{k=c+1}^{M} (k-c) \binom{M}{k} p^k (1-p)^{M-k} \tag{12}$$

This formula can be used to estimate how many talkers can be handled on a particular channel for a given required upper bound on fractional packet loss.

It can be commented that the above formula for φ applies without change to the fraction of speech lost in a TASI system (analog or digital) which shares c circuits among M talkers. The detailed effects on particular users will be different in the TASI system, since a talkspurt finding no circuit available will stay frozen out until a circuit becomes available. But the talker activity model used above applies equally to a TASI system, and the fractional speech loss could be derived as above by letting T_p become the minimum time interval needed to make a talkspurt/silence decision in the speech activity detector.

REFERENCES

1. J. W. Forgie and A. G. Nemeth, "An Efficient Packetized Voice/Data Network using Statistical Flow Control," to be published in Proceedings of 1977 International Conference on Communications.
2. K. Bullington and J. M. Fraser, "Engineering Aspects of TASI," Bell Syst. Tech. J. 38, 353-364 (March 1959).
3. A. G. Nemeth, "Behavior of a Link in a PVC Network," Technical Note 1976-45, Lincoln Laboratory, M.I.T. (November 1976), DDC AD-A036370.
4. L. Kleinrock, Queueing Systems, Volume 1: Theory (Wiley, New York, 1976), pp. 107-108.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-77-187	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Information Processing Techniques Program Volume II: Communications-Adaptive Internetting		5. TYPE OF REPORT & PERIOD COVERED Semiannual Technical Summary 1 July 1976 – 31 March 1977
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR*(s)* Bernard Gold		8. CONTRACT OR GRANT NUMBER*(s)* F19628-76-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, M.I.T. P.O. Box 73 Lexington, MA 02173		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ARPA Order 2929 Program Element No. 62708E Project No. 7T10
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, VA 22209		12. REPORT DATE 31 March 1977
		13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)* Electronic Systems Division Hanscom AFB Bedford, MA 01731		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the obstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

Volume I is ESD-TR-77-81.

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

information processing techniques
communications-adaptive internetting
digital voice communications
packet speech
network speech terminal
C^2 links
time-varying communications

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This report describes work performed on the Communications-Adaptive Internetting program sponsored by the Information Processing Techniques Office of the Defense Advanced Research Projects Agency. The Semiannual Technical Summary preceding this one was entitled "Airborne Command and Control" and covered the period ending 30 June 1976. In order to come into compliance with the current fiscal year schedule, this report has been delayed three months and covers the period 1 July 1976 through 31 March 1977.